

5

including the additional value for the additional parameter in the third-party invocation request. In some of those implementations, selecting the particular third-party agent from the group of third-party agents includes: selecting the particular third-party agent based on the value and the additional value conforming to corresponding values defined, for the third-party agent, for the mandatory parameter and the additional parameter. In some versions of those implementations, selecting the particular third-party agent from the group of third-party agents further includes: eliminating an additional third-party agent from the selecting based on the value and the additional value not conforming to corresponding values defined, for the additional third-party agent, for the mandatory parameter and the additional parameter. The additional third-party agent is of the group of third-party agents. Selecting the particular third-party agent may in some of those implementations occur without providing the user with output that explicitly solicits the user to select between the third-party agent and one or more additional third-party agents of the group.

In some implementations, a method performed by one or more processors is provided and includes, as part of a dynamic dialog between a user and an automated assistant implemented by one or more of the processors: receiving natural language input of a user that is directed to the automated assistant as part of the dynamic dialog, the natural language input being free-form and being based on user interface input provided by the user via a user interface input device; selecting an intended action based on determining that the natural language input indicates the intended action, the intended action being selected from a group of potential intended actions that span multiple service sectors; identifying a group of third-party agents that can each perform the intended action indicated by the natural language input and that can each perform the intended action; generating a prompt that is based on a parameter that is stored in association with the intended action and that is unspecified in the natural language input, where the prompt does not explicitly identify any of the third-party agents of the group; providing the prompt as part of the dynamic dialog and as a reply to the natural language input; receiving additional natural language input provided by the user as part of the dynamic dialog in response to the prompt; determining a value for the parameter based on the additional natural language input; eliminating a third-party agent from the group of third-party agents to form a subgroup, the eliminating based on the value not conforming to a corresponding value defined for the parameter for the third-party agent; selecting a particular third-party agent from the subgroup; and transmitting a third-party invocation request that includes the value for the parameter. The transmitting is to only the selected particular third-party agent via one or more network interfaces.

These and other implementations of technology disclosed herein may optionally include one or more of the following features.

In some implementations, the method further includes: receiving, via one or more of the network interfaces, responsive content from the third-party agent in response to the transmitting the third-party invocation request; and providing, for presentation to the user as part of the dynamic dialog, output that is based on the responsive content. In some of those implementations, the method further includes: receiving further input provided by the user in response to the output that is based on the responsive content of the third-party agent; and transmitting, to the third-party agent via one or more of the network interfaces, content that is

6

based on the further input. In some of those implementations, the further input is voice input and the method further includes: converting the voice input to text and including at least some of the text in the content.

In some implementations, a method performed by one or more processors is provided and includes: receiving, by a local agent over one or more network interfaces, first natural language input transmitted by a client device of a user. The first natural language input is free-form and is directed to an automated assistant interface implemented by the client device. The method further includes determining, by the local agent, that the first natural language input indicates a first intended action that is directed to a first service sector. The method further includes, by the local agent via one or more of the network interfaces: interacting with the user via the automated assistant interface of the client device to resolve values for all first mandatory parameters stored in association with the first intended action. The method further includes selecting, by the local agent, a particular third-party agent from a group of third-party agents that can each perform the intended action indicated by the first natural language input. The method further includes, by the local agent subsequent to resolving the values for the first mandatory parameters: invoking the third-party agent based on the resolved values for the first mandatory parameters; and facilitating interactions between the client device and the third-party agent that occur via the automated assistant interface, and that are steered by the third-party agent.

In some implementations, the method further includes: receiving, by the local agent over one or more of the network interfaces, second natural language input from the client device of a user, the second natural language input being free-form and being directed to the automated assistant interface; determining, by the local agent, that the second free-form natural language input indicates a second intended action that is directed to a second service sector, the second service sector being disparate from the first service sector; by the local agent via one or more of the network interfaces, interacting with the user via the automated assistant interface of the client device to resolve second values for all second mandatory parameters stored in association with the second intended action; selecting, by the local agent, an additional particular third-party agent from a group of additional third-party agents that can each perform the second intended action indicated by the natural language input, the group of additional third-party agents being disparate from the group of third-party agents; and by the local agent subsequent to resolving the second values for the second mandatory parameters: invoking the additional third-party agent based on the resolved second values for the second mandatory parameters, and facilitating interactions between the client device and the additional third-party agent that occur via the automated assistant interface, and that are steered by the additional third-party agent.

In some implementations, a method performed by one or more processors is provided and includes: receiving an initial voice input provided by a user via a client device; performing a voice to text conversion to convert the initial voice input to initial text; determining an intended action based on the initial text; identifying a mandatory parameter that is stored as mandatory for the intended action; determining that the initial text lacks specification of any value for the mandatory parameter; in response to determining that the initial text lacks specification of any value for the mandatory parameter: generating a natural language prompt based on the mandatory parameter, and providing the natural language prompt as a reply to the initial voice input, the